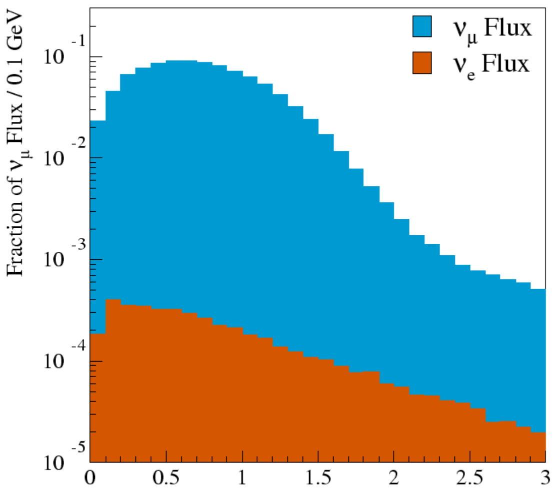
Systematic errors of MiniBooNE

Kendall Mahn, Columbia University
for the MiniBooNE collaboration

NuFact07

Oscillation Analysis

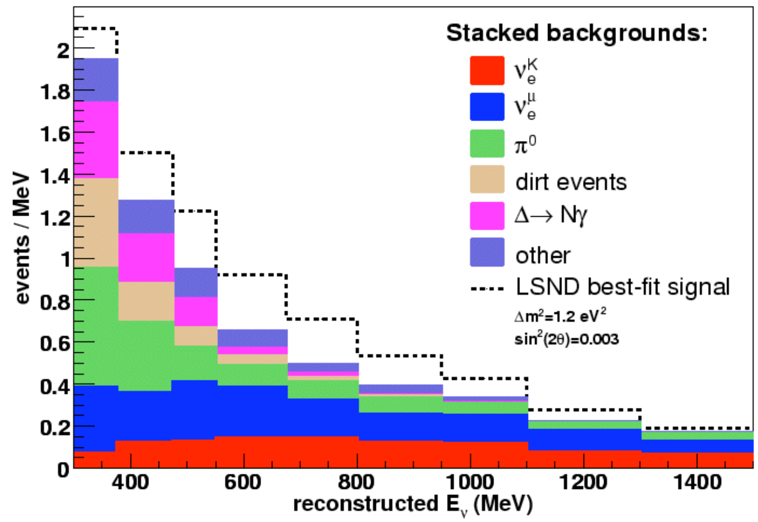


ν_μ

E_ν(QE)

0.5% intrinsic ν_e

ν_e selection cuts



Signal

E_ν(QE)

(Δm^2=1.2eV^2, $\sin^2 2\theta$=0.003)

Background

- misidentified ν_μ (mainly π^0s)
- ν_e from μ^+ decay
- ν_e from K^+, K^0 decay
- $\Delta \Rightarrow N\gamma$
- Out of tank events ('dirt')

<u>Do the ν_μ oscillate into ν_e ?</u>

- Produce ν_μ
- Select ν_e
- Observe an excess or not? Check if the excess is consistent with oscillation

Strategy

- Modern neutrino beam experiments use a ‘near to far’ ratio to observe oscillations
  - Directly compare initial neutrino beam to final neutrino beam
  - This causes many systematic errors to cancel between the two samples, such as flux and cross sections
- MiniBooNE has no near detector, but we can still use measurements to constrain backgrounds
- Two complementary analysis approaches address constraints
  - Use a data sample to correct a background
  - Fit two samples simultaneously to reduce the size of the errors

Constraints with MiniBooNE data

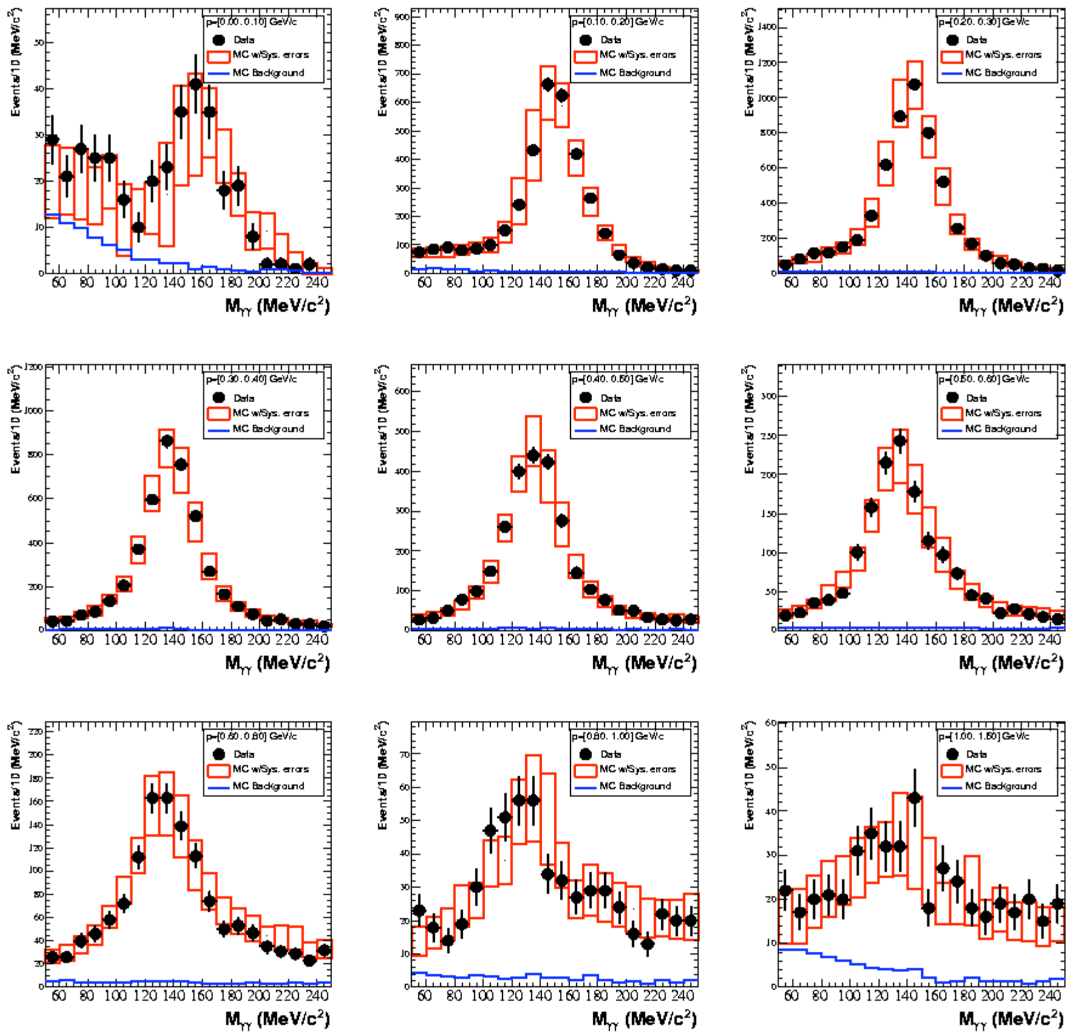
Use a data sample to correct a background

- We must learn about the signal region without looking at it (blind analysis)
- Measure pure or enhanced samples of a given background; rate measurements circumvent flux, cross section errors
- Infer the shape and normalization of the background in the signal region
- Examples: NC π^0 misIDs, out of tank events, ν_e from μ decay

Constraint Example: NC π^0s

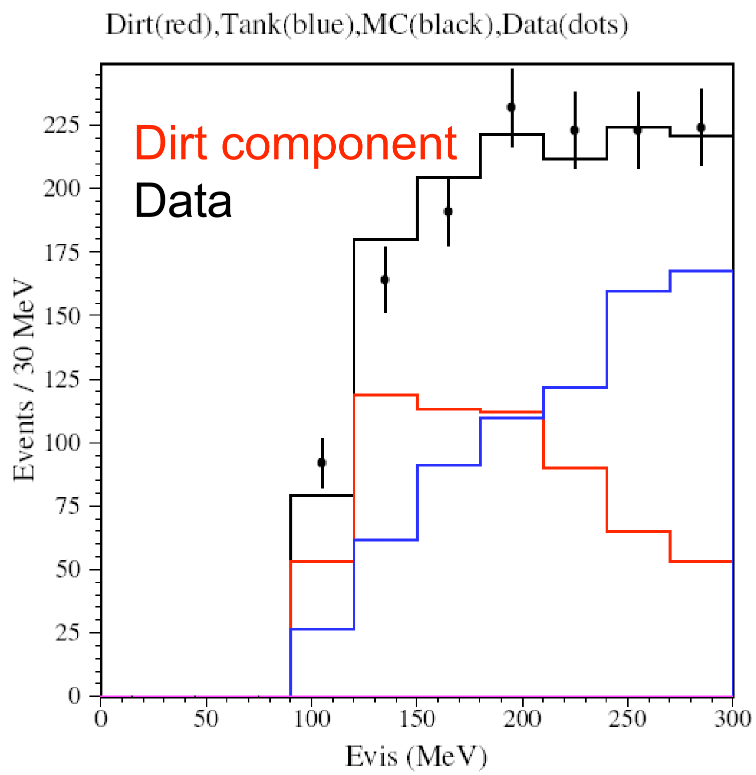
- Measure π^0s in MiniBooNE
  very pure (~90%) sample
- Compare the observed π^0 rate to the MC as a function of π^0 momentum, and make a correction factor
- Reweight the misidentified π^0s in the ν_e sample based on their momentum by this correction factor
- Can also correct radiative events $\Delta \rightarrow N + \gamma$

$M_{\gamma\gamma}$ Mass Distribution for Various $p_{\pi 0}$ Momentum Bins



Constraint Example: Out of tank events

- Events from interactions in the surrounding rock produce photons which pass the veto and give events within the inner tank (so called "dirt") events
- Create a sample of enhanced dirt events
  - in time with beam, minimal veto activity, 1 subevent, not decay electron
  - low energy, high radius
- Checks prediction spatial distribution, energy spectrum of these events; sets the normalization for dirt events in the ν_e sample



visible energy (MeV)

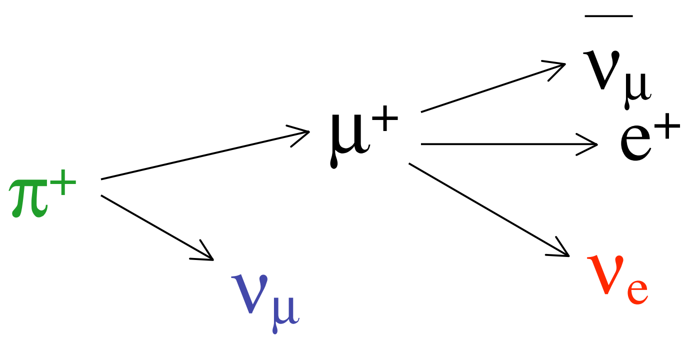
Constraint Example: ν_e from μ^+

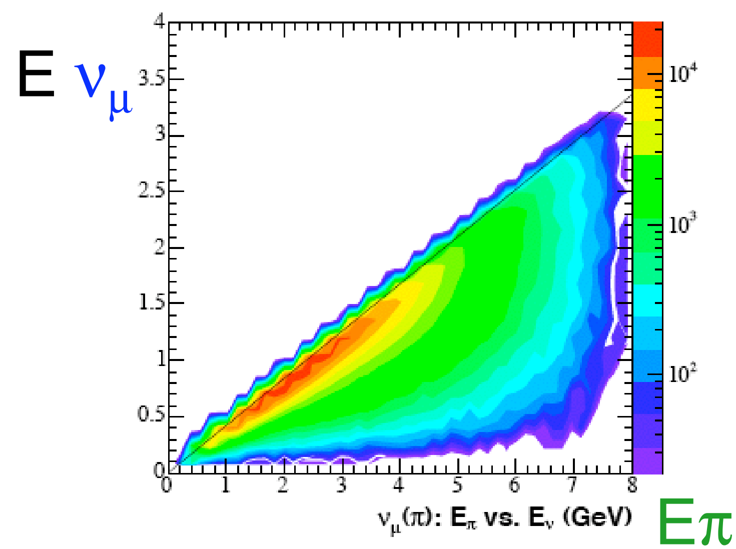
Without employing a link between ν_e and ν_μ , ν_e from μ^+ would have flux, cross section, detector uncertainties

However, for each ν_e produced from a μ^+, there was a corresponding ν_μ and we observe that ν_μ spectrum

This is true here because the pion decay is very forward

Therefore, we know that some combination of cross sections, flux, etc errors are excluded by our own data, and so the error is reduced





Constraints in action

Two methods to include ν_μ information into the ν_e analysis:

- Reweight the ν_e based on the observed ν_μ spectrum, and then fit the ν_e s for oscillation (used in likelihood analysis)
- Fit simultaneously the ν_μ and ν_e energy spectrums (used in boosted decision tree analysis)
  - ν_μ provide information to constrain errors, ν_e provide information for oscillation parameters

Fit Mechanics

To fit data d to some prediction p, form a $\chi 2$:

$$\chi^2 = \sum_{i,j=1}^{bins} \Delta_i M_{ij}^{-1} \Delta_j$$

where Δ = (d-p) in each energy bin i or j. 2 parameter mixing scenario included in p

$(M_{ij})^{-1}$ is the inverse of the error matrix

Systematic (and statistical) uncertainties in M_{ij} matrix

If only it were this simple...

If M_{ij} were just statistics, it would have values along the diagonals, and zero elsewhere. This matrix has no correlations, as each bin contributes to the $\chi 2$ only as the square of itself.

$$M_{ij} = \begin{matrix} N_1 & 0 & 0 & 0 & 0 \\ 0 & N_2 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & N_k & 0 \\ 0 & 0 & 0 & 0 & N_k \end{matrix}$$

To construct this matrix for any set of uncertainties α, one would measure each α and sum the square of the error in each bin:

$$M_{ij} = \sum_{\alpha=1}^{systematics} \sigma^2{}_{ij}(\alpha)$$

$$M_{ij} = \begin{matrix} (\sigma^2{}_1+\sigma^2{}_2+\ldots+\sigma^2{}_\alpha)_1 & 0 & 0 & 0 & 0 \\ 0 & (\sigma^2{}_1+\sigma^2{}_2+\ldots+\sigma^2{}_\alpha)_2 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & (\sigma^2{}_1+\sigma^2{}_2+\ldots+\sigma^2{}_\alpha)_{k-1} & 0 \\ 0 & 0 & 0 & 0 & (\sigma^2{}_1+\sigma^2{}_2+\ldots+\sigma^2{}_\alpha)_k \end{matrix}$$

... now to reality

Consider a single source of error, but now with correlations:

$$M_{ij} = \begin{pmatrix} \sigma^2_{11} & \rho_{21}\sigma_2\sigma_1 & \cdots & \rho_{k1}\sigma_k\sigma_1 \\ \rho_{12}\sigma_1\sigma_2 & \sigma^2_{22} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \rho_{kk-1}\sigma_k\sigma_{k-1} \\ \rho_{1k}\sigma_1\sigma_k & \vdots & \rho_{k-1k}\sigma_{k-1}\sigma_k & \sigma^2_{kk} \end{pmatrix}$$

Now, bin i is related to bin j by $\rho_{ij}\sigma_i\sigma_j$

Fit Mechanics

Do a combined oscillation fit to the observed ν_μ and ν_e energy distribution

$$\chi^2 = \begin{pmatrix} \Delta_i^{\nu_e} & \Delta_i^{\nu_\mu} \end{pmatrix} \begin{pmatrix} M_{ij}^{e,e} & M_{ij}^{e,\mu} \\ M_{ij}^{\mu,e} & M_{ij}^{\mu,\mu} \end{pmatrix}^{-1} \begin{pmatrix} \Delta_j^{\nu_e} \\ \Delta_j^{\nu_\mu} \end{pmatrix}$$

where $\Delta_i^{\nu_e} = \mathrm{Data}_i^{\nu_e} - \mathrm{Pred}_i^{\nu_e}(\Delta m^2, \sin^2 2\theta)$ and $\Delta_i^{\nu_\mu} = \mathrm{Data}_i^{\nu_\mu} - \mathrm{Pred}_i^{\nu_\mu}$

Note this $\chi2$ includes ν_μ sample and ν_e sample bins, and a 2 parameter oscillation scenario. M_{ij} has 4 distinct sections: ν_e / ν_e bin terms, ν_μ / ν_μ bin terms, and cross terms which mix ν_μ and ν_e

$$M_{ij} = \begin{pmatrix} \nu_e & \nu_e / \nu_\mu \\ \nu_\mu / \nu_e & \nu_\mu \end{pmatrix}$$

How this helps: 2x2 case

Take just 1 ν_μ , ν_e bin:

$$M_{ij} = \begin{pmatrix} Ne + \sigma^2 e & \rho\sigma e\sigma\mu \\ \rho\sigma\mu\sigma e & N\mu + \sigma^2\mu \end{pmatrix}$$

Invert, and multiply by (Δ_e Δ_μ) Δ= data-prediction(signal). The χ2 minimizes for signal of:

$$signal = \Delta e\left(1 - \frac{\rho}{(N\mu/\sigma\mu + 1)}\frac{\Delta\mu/\sigma\mu}{\Delta e/\sigma e}\right)$$

with an uncertainty of:

$$\sigma^2{}_{signal} = Ne + \sigma e^2\left(1 - \frac{\rho^2}{(N\mu/\sigma\mu + 1)}\right)$$

With ρ approaching 1 (high correlation) and small statistical error for ν_μ:

$$signal = \Delta e\left(1 - \frac{\Delta\mu/\sigma\mu}{\Delta e/\sigma e}\right) \pm Ne$$

or the error on the signal is limited by the statistical error, not systematic error of the ν_e sample

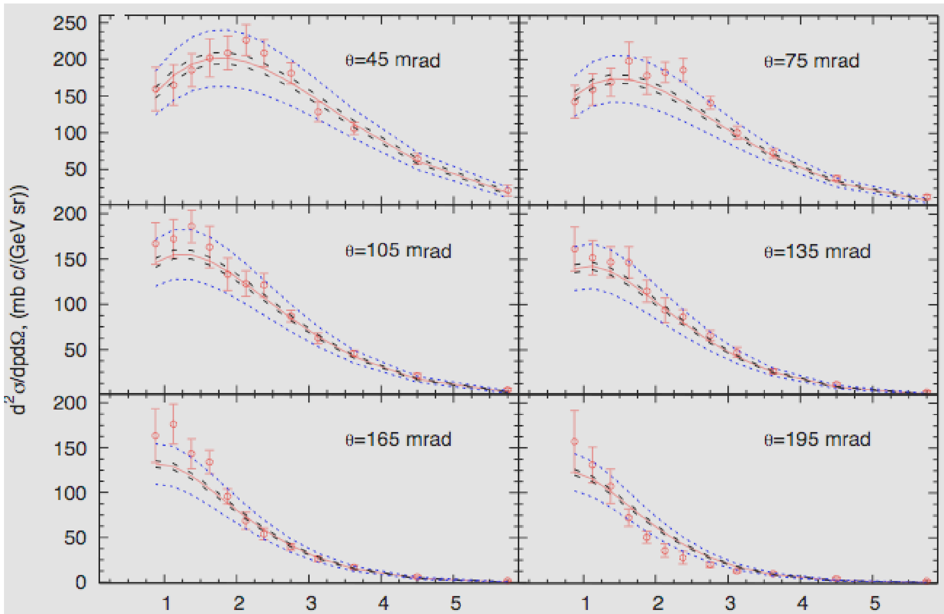
Building an error matrix

For each error, build a error matrix, and then sum for final error matrix

Flux from π^+/μ^+ decay	$M_{ij}(\pi+)$
Flux from K^+ decay	$+M_{ij}(K+)$
Flux from K^0 decay	$+M_{ij}(K0)$
Target/Beam model	$+M_{ij}(tar\,/\,beam)$
ν cross section	$+M_{ij}(x\sec)$
NC π^0 yield	$+M_{ij}(NC\pi 0)$
Out of tank events	$+M_{ij}(dirt)$
Optical Model	$+M_{ij}(OM)$
DAQ electronics model	$+M_{ij}(DAQ)$
	$= M_{ij}(total)$

Building an error matrix: π+ production

Take existing data (HARP 8.9 GeV/c pBe π^+ production data) and fit it to a parameterization (Sanford-Wang)



The fit gives the 9 parameters c_i and their errors

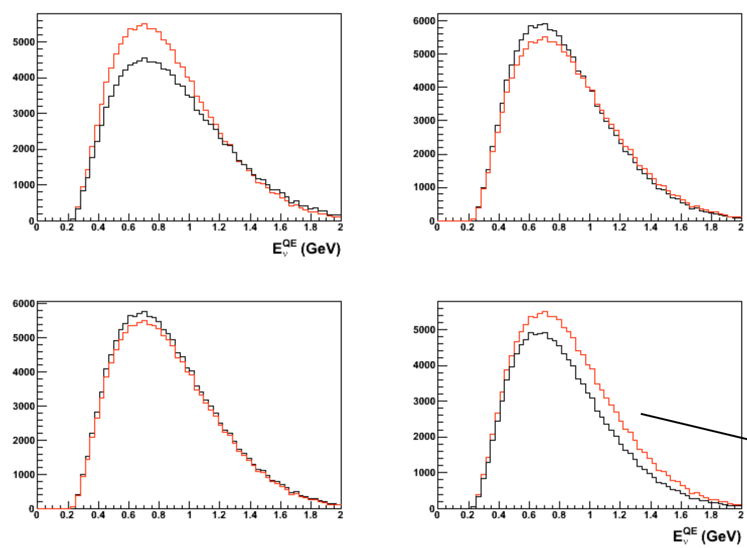
The parameterization provides correlations amongst the c_i (covariance matrix)

+

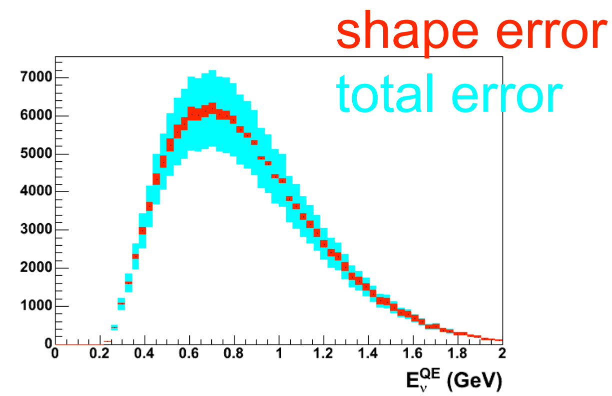
$$\frac{d^2\sigma(p+A\to\pi^++X)}{dp\, d\Omega}(p,\theta) = c_1 p^{c_2}(c_9 - p/p_{beam}) \exp[-c_3\,(p^{c_4}/p_{beam}{}^{c_5}) - c_6\theta(p - c_7 p_{beam} \cos^{c_8}\theta)\,]$$

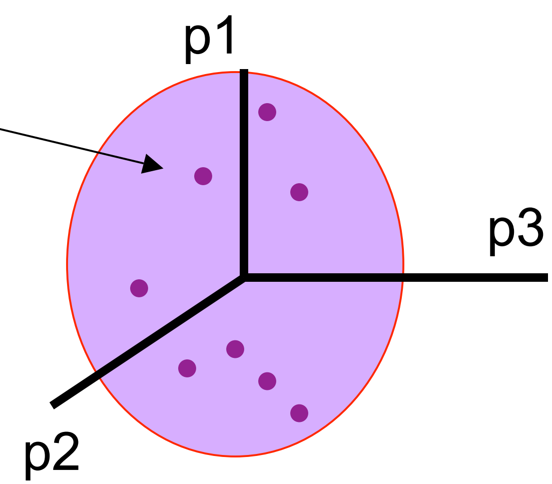
Building an error matrix: π+ production

Throw the c_i according to their covariance matrix and within their errors many many times...



+ ... =



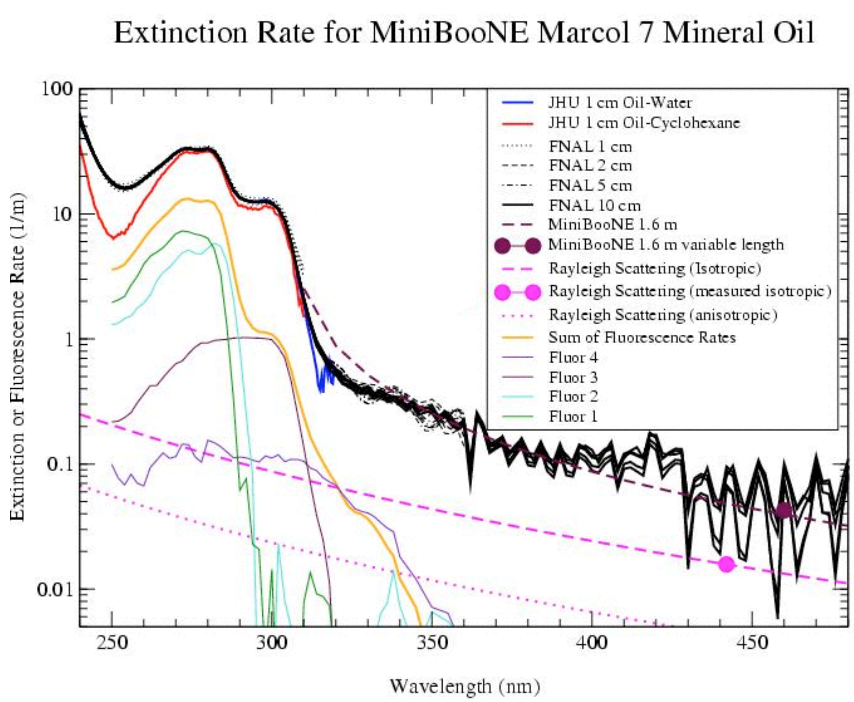


The error matrix:

$$M_{ij}^{\pi+prod} = \frac{1}{throws} \sum_{k=1}^{throws} (N_{cv} - N_k)_i (N_{cv} - N_k)_j$$

Building an error matrix: light propagation in detector

For the optical model, use a combination of external and internal measurements to produce the covariance matrix

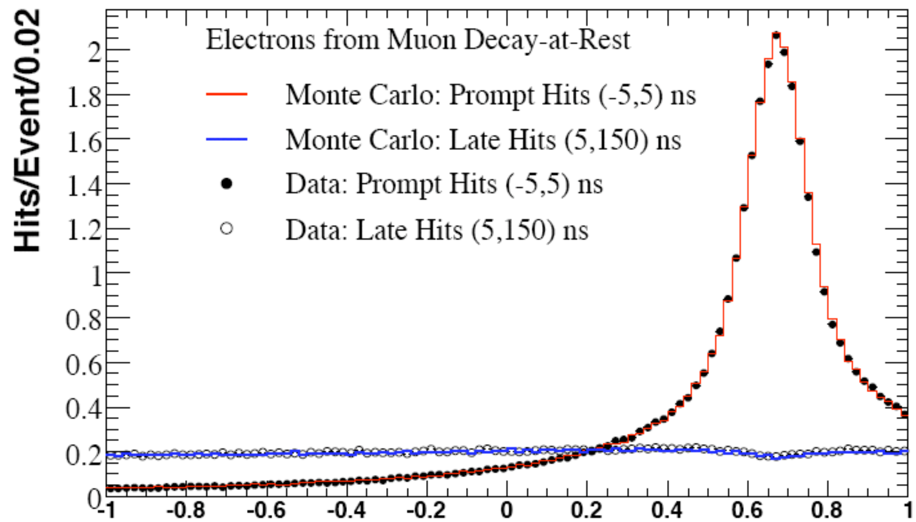


Use measurements of oil, PMTs to decide model's (39!) parameters and initial errors

- Scintillation from p beam (IUCF)
- Scintillation from cosmic μ (Cincinnati)
- Fluorescence Spectroscopy (FNAL)
- Time resolved spectroscopy (JHU, Princeton)
- Attenuation (Cincinnati)

Building an error matrix: light propagation in detector

Create different 'universes' with the parameters varied within errors

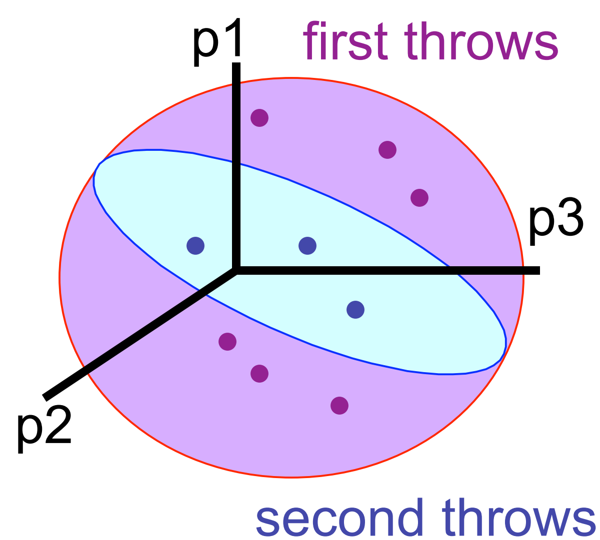


Compare them to muon decay electron (Michel) sample variables, such as time, charge, hit topology

Keep universes which have a good $\chi 2$ as compared to data

This restricted space defines the parameters and correlations. Draw from the new space, and build an error matrix:

$$M_{ij}^{OM} = \frac{1}{universes} \sum_{k=1}^{universes} (N_{cv} - N_k)_i (N_{cv} - N_k)_j$$



Building an error matrix:
light propagation in detector

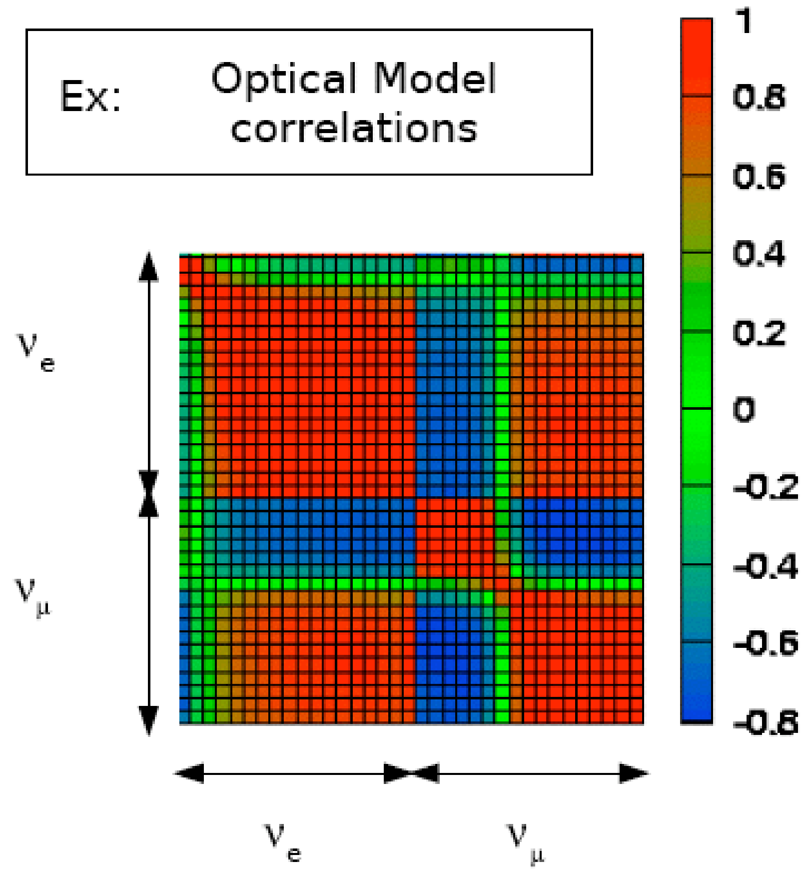
Example: Optical model final error matrix

highly correlated

highly anticorrelated

not correlated

$$M_{ij} = \begin{pmatrix} \nu_e & \nu_e / \nu_\mu \\ \nu_\mu / \nu_e & \nu_\mu \end{pmatrix}$$



Error 'budget'

source of uncertainty on ν_e background	TBL/BDT % error	constrained by MB data?	Reduced by relating ν_μ to ν_e
Flux from π^+/μ^+ decay	6.2 / 4.3	Y	Y
Flux from K^+ decay	3.3 /1.0	Y	Y
Flux from K^0 decay	1.5 / 0.4	Y	Y
Target/Beam model	2.8 / 1.3	Y	Y
ν cross section	12.3 / 10.5	Y	Y
NC π^0 yield	1.8 / 1.5	Y	
Out of tank events	0.8 / 3.4	Y	
Optical Model	6.1 / 10.5	Y	Y
DAQ electronics model	7.5 / 10.8	Y	

All of our errors are highly correlated, but here are the diagonal errors

* shows errors before ν_e / ν_μ constraint is applied

Summary

Many oscillation experiments employ a near to far ratio to reduce their systematic errors; MiniBooNE uses a 'ν_e / ν_μ' ratio to reduce errors

MiniBooNE constrains all backgrounds with data samples

The error formalism includes all correlations between ν_μ and ν_e, which are then exploited in the final fit

- ν_μ small statistical error lowers the ν_e effective systematic error